DOCUMENT RESUME

ED 024 156 EA 001 890

By- Stufflebeam, Daniel L.

Evaluation Under Title 1 of the Elementary and Secondary Educational Act of 1965.

Pub Date 24 Jan 66

Note- 23p.; Paper presented at Evaluation Conference sponsored by the Michigan St. Dep. of Educ. (East Lansing, Mich., January 24, 1966).

EDRS Price MF-$0.25 HC-$1.25

Descriptors- Data Collection, *Decision Making, Educational Change, *Evaluation, *Federal Legislation, *Program Evaluation, State Departments of Education

Identifiers- Elementary and Secondary Education Act, Title 1

Evaluation is discusses as a primary element in the change process. Defined as a procedure for obtaining data to make decisions, evaluation is regarded as a recurring determination of the extent to which objectives have been reached, within a four-stage sequence of research, development, diffusion, and adoption. A conceptual scheme for viewing the role of evaluation is explained and distinctions are made between evaluation procedures applicable to decision making on Federal and State administrative levels and procedures applicable to decision making on the local level. Six recommendations are made for the implementation of evaluation of projects developed under the ESEA Act of 1965. (JK)

EDO 24150

EVALUATION UNDER TITLE I OF THE ELEMENTARY AND SECONDARY EDUCATIONAL ACT OF 1965[1]

Daniel L. Stufflebeam
Director, Evaluation Center
The Ohio State University

You have asked me to talk about some of the problems and possible solutions associated with the evaluation of projects to be conducted under Title I of the Education Act.[2] I am intensely interested in this topic and am currently devoting most of my efforts to it. The more I learn about the Title I program, the more I am impressed by the wide range of difficult evaluation questions which educators are asking and by the lack of practical answers that are available. Hopefully, during this conference we can tackle some of these tough questions and move toward reaching some viable, operational-level solutions.

Each of you, I am certain, is acutely aware of the urgent need to do this. Under Title I alone, over one billion dollars is authorized for fiscal year 1966. While this is an unprecedented opportunity to improve education, it also brings a challenge; the schools must provide for effective evaluations. That is, they must present objective data which demonstrate the effectiveness of their new plans and programs.

---

[1]Address at Evaluation Conference sponsored by the Michigan State Department of Education, East Lansing, January 24, 1966.

[2]Public Law 89-10, The Elementary and Secondary Education Act of 1965.

EA 001 890

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

The need for effective evaluations takes on increased importance when one considers the implications of such evaluations. At the Federal level, allocations for succeeding years will be largely dependent upon evaluation results for programs conducted during these early years. At the local level, decisions about retaining, abandoning, or revising programs can be made more judiciously if sound, objective data are available. It is obvious that solutions to the evaluation problem must be found immediately. Further, such solutions must be tailored to the intent and characteristics of the Education Act.

NATURE OF THE EDUCATION ACT

The purpose of the Education Act is to meet the special educational needs of deprived children. Somewhat more specifically, it is to increase the educational attainment, experiences, and opportunities of these children. The program is national in scope, design, support and broad control. It is to be coordinated and specifically controlled at the state level and to be implemented in the local school districts. The over-riding, long-range purpose, as I see it, is to promote planned change in education on a continuing basis, that is, to close the gap between theory and practice.

To achieve this broad purpose, many changes obviously will be required across the face of education. New methods and materials of many kinds will have to be developed. These will have to be publicized and demonstrated to the schools which need such innovations. The local schools will have to adopt and adapt those innovations which are best fitted to their needs. Thus, a wide range of activities are appropriate under the Act. Further, these will need to be coordinated into a program

of planned change in education, if the impact of the Title I program is to have far-reaching implications for meeting the needs of disadvantaged children, and if wasteful duplications of effort are to be avoided.

A Scheme for Classifying Education Act Projects

David L. Clark and Egon G. Guba[3] have proposed a taxonomy for classifying the many activities that may be sanctioned under the Act. Originally, they developed this taxonomy to cover the range of activities necessary for producing planned change in education. I do not wish to enter a discussion of change process research and theory per se, but I do want, from time to time this morning, to refer to the Clark-Guba taxonomy in order to illustrate the process and function of evaluation within the total framework of change and innovation. I should, therefore, now like to give you a brief introduction to this taxonomy.

Figure 1 contains a modified version of the Clark-Guba taxonomy. (Any criticisms that you may wish to make of this taxonomy should be directed toward me, since I have adapted it to fit my purposes. But, any credit should go to Drs. Guba and Clark since they master-minded the development of the original scheme).

You will note that I have converted this taxonomy into the form of a flow diagram, the flow being from top to bottom. The change process as depicted here includes the activities of research, development, diffusion, adoption, and evaluation. Research is depicted by a triangle, development

---

[3] David L. Clark and Egon G. Guba, An Examination of Potential Change Roles in Education. Paper presented before the Seminar on Innovation in Planning School Curricula, October, 1965.

A PROCESS CHART DEPICTING THE ROLE OF EVALUATION IN THE CHANGE PROCESS

	AGENCY	OBJECTIVE	PROCESS	CRITERIA	RELATION TO CHANGE
RESEARCH	UNIVERSITY	TO ADVANCE KNOWLEDGE	RESEARCH → EVALUATION	VALIDITY (INTERNAL AND EXTERNAL)	PROVIDES BASIS FOR INVENTION
DEVELOPMENT	UNIVERSITIES, LOCAL SCHOOLS, INDUSTRY, AND RESEARCH AND DEVELOPMENT INSTITUTIONS	TO FORMULATE A NEW SOLUTION TO AN OPERATING PROBLEM OR TO A CLASS OF OPERATING PROBLEMS, I.E., TO INNOVATE	INVENTION → EVALUATION	FACE VALIDITY (APPROPRIATENESS); ESTIMATED VIABILITY; IMPACT (RELATIVE CONTRIBUTION)	PRODUCES THE INVENTION
		TO DRAFT A PLAN FOR CONSTRUCTING THE INNOVATION, I.E., TO CONSTRUCT THE BLUEPRINT.	DESIGN → EVALUATION	FEASIBILITY (PRODUCTION AND UTILIZATION); TRACTABILITY (EASE OF MANAGING, CONTROLLING, AND INSTRUCTING IN THE USE OF)	ENGINEERS THE INVENTION TO FIT THE CHARACTERISTICS OF THE TARGET SITUATION
		TO BUILD THE COMPONENTS	CONSTRUCTION → EVALUATION	DESIGN SPECIFICATIONS; INDIVIDUAL PERFORMANCE	PRODUCES THE COMPONENTS NECESSARY FOR IMPLEMENTING THE DESIGN
		TO INTEGRATE THE COMPONENTS INTO AN OPERATING SYSTEM	ASSEMBLY → EVALUATION	TOTAL PERFORMANCE; VIABILITY; EFFICIENCY	PRODUCES THE COORDINATED OPERATING SYSTEM
DIFFUSION	GOVERNMENT AND UNIVERSITIES	TO CREATE WIDESPREAD AWARENESS OF THE INVENTION AMONG PRACTITIONERS, I.E., TO INFORM	DISSEMINATION → EVALUATION	INTELLIGIBILITY; FIDELITY; PERVASIVENESS; IMPACT (EXTENT TO WHICH IT AFFECTS KEY TARGETS)	INFORMS ABOUT THE INVENTION
		TO AFFORD AN OPPORTUNITY TO EXAMINE AND ASSESS OPERATING QUALITIES OF THE INVENTION, I.E., TO BUILD CONVICTION	DEMONSTRATION → EVALUATION	CREDIBILITY; CONVENIENCE; EVIDENTIAL ASSESSMENT	BUILDS CONVICTION ABOUT THE INVENTION
ADOPTION	SCHOOLS	TO BUILD FAMILIARITY WITH THE INVENTION AND PROVIDE A BASIS FOR ASSESSING THE QUALITY, VALUE, FIT, AND UTILITY OF THE INVENTION IN A PARTICULAR INSTITUTION, I.E., TO TEST	TRIAL → EVALUATION	ADAPTABILITY; FEASIBILITY; ACTION	TRIES OUT THE INVENTION IN THE CONTEXT OF A PARTICULAR SITUATION
		TO FIT THE CHARACTERISTICS OF THE INVENTION TO THE CHARACTERISTICS OF THE ADOPTING INSTITUTION, I.E., TO OPERATIONALIZE	INSTALLATION → EVALUATION	EFFECTIVENESS; EFFICIENCY	OPERATIONALIZES THE INVENTION FOR USE IN A SPECIFIC INSTITUTION
		TO ASSIMILATE THE INVENTION AS AN INTEGRAL AND ACCEPTED COMPONENT OF THE SYSTEM, I.E., TO ESTABLISH	INSTITUTIONALIZATION → EVALUATION	CONTINUITY; VALUATION; SUPPORT	ESTABLISHES THE INVENTION AS A PART OF AN ONGOING PROGRAM; CONVERTS IT TO A "NON-INNOVATION"

*Based upon "A Classification Scheme of Processes Related to and Necessary for Change in Education," by David L. Clark and Egon G. Guba.

by convex hexagons, diffusion by concave hexagons, adoption by rectangles, and evaluation by diamonds. Each activity (with the exception of evaluation) is classified, from left to right, by the agencies which may be expected to be primarily involved in that activity, by the objective associated with the activity, by some of the criteria associated with the objective, and by the relation of the activity to the change process. I shall be concerned primarily with the objectives and criteria associated with each activity and with the overall relation of evaluation activities to the change process. Before discussing this chart further, I now want to move to a consideration of the nature of evaluation in the Title I program.

EVALUATION DEFINED

The guidelines for Title I define evaluation as the process of determining the extent to which objectives have been reached. To evaluate, it is therefore necessary to know what the objectives are and to specify criteria for determining the relation of the outcomes to the objectives. Since different kinds of activities have different objectives, different kinds of evaluations may be distinguished.

From Figure 1 you will note that an evaluation activity immediately follows each of the other change activities. This is intended to depict that evaluation is appropriate at each stage of the change process and that different kinds of evaluations may be distinguished for different levels of objectives and criteria.

For example, in test development, evaluations of test plans differ from evaluations of test items. The objective associated with developing a test plan is to <u>design</u> or <u>blueprint</u> the test to be developed. The

criteria for evaluating a test plan are content validity, feasibility, and estimated reliability. The usual method of evaluating a test plan is to secure expert judgment, e.g., from curriculum and measurement specialists. A method quite different from this one would be employed in evaluating the actual test questions once they were developed. The objective in developing test questions is to construct questions which meet the design specifications. The criteria that are appropriate for evaluating test questions are design specifications (i.e., do the questions cover the specified content and objective areas?) and individual performance (i.e., do the questions possess concurrent validity and reliability?). The method usually employed in evaluating test questions is to administer them and certain criteria instruments to a sample of students and then to perform item analyses.

We have seen from this example that different approaches to evaluation were required, due to different objectives and criteria for the two activities. In the first case, expert judgment was employed; in the second, statistical analyses were used. This example could be extended to other cases and other phases of the diagram, but in the interest of time I shall not attempt this. I think the point is clear: evaluations may differ widely as a function of objectives and criteria. I should now like to move to a consideration of the role of evaluation under Title I.

THE ROLE OF EVALUATION UNDER TITLE I

As I see it, the purpose of evaluation in the Title I program is to assist in making decisions. Based on evaluative data, decisions will be made to continue, discontinue, or modify projects and practices.

Further, such decisions will be made at three administrative levels: Federal, state and local. Evaluations must, therefore, be designed to provide the data that will be required for making such decisions.

Returning to Figure 1 once again, you will note that arrows lead from each evaluation activity to both the preceding change activity and the succeeding change activity. This is intended to depict the decision function of evaluation. At each stage in the change process, outcomes are compared with objectives on the basis of established criteria. If the criteria have been met, the process flow continues to the next activity; if not, the flow reverts to a prior activity in order that corrections can be made. I realize that reality will not often conform to this neat diagram, that in the change process many of the depicted change phases may often be skipped, and that the alternative of "scrubbing" a mission (to borrow a phrase from the space industry) is always present at any phase of the change process, for any particular activity. Although Figure 1 does not adequately account for all such alternatives I think it does depict the general function of evaluation. This function, as I see it, is to insure sound, systematic progress toward putting theory into practice and to serve as a means for identifying problems and taking corrective action as each failure is encountered. A further function, not so well depicted by this chart, is an information feedback function. Evaluations can continually provide information pertinent to the conduct of other projects. For example, evaluative information about a demonstration project might well be useful to a researcher who is studying educational change.

It is important to note that decisions about Title I practices will be made at Federal, state, and local levels, while the bulk of basic data for making these decisions will be collected at the local level alone. Thus the evaluation designs of local schools must be consistent with state and Federal requirements for evaluation. Obviously, then, Federal and state objectives and criteria must be specified and, further, specific evaluation designs must be developed for each of these levels. Only after this has been accomplished can local schools prepare evaluation designs which will provide the required data.

I should now like to suggest a scheme for viewing in greater detail the decision process as it may occur in the Title I program.

A CONCEPTUAL SCHEME FOR VIEWING THE DECISION PROCESS

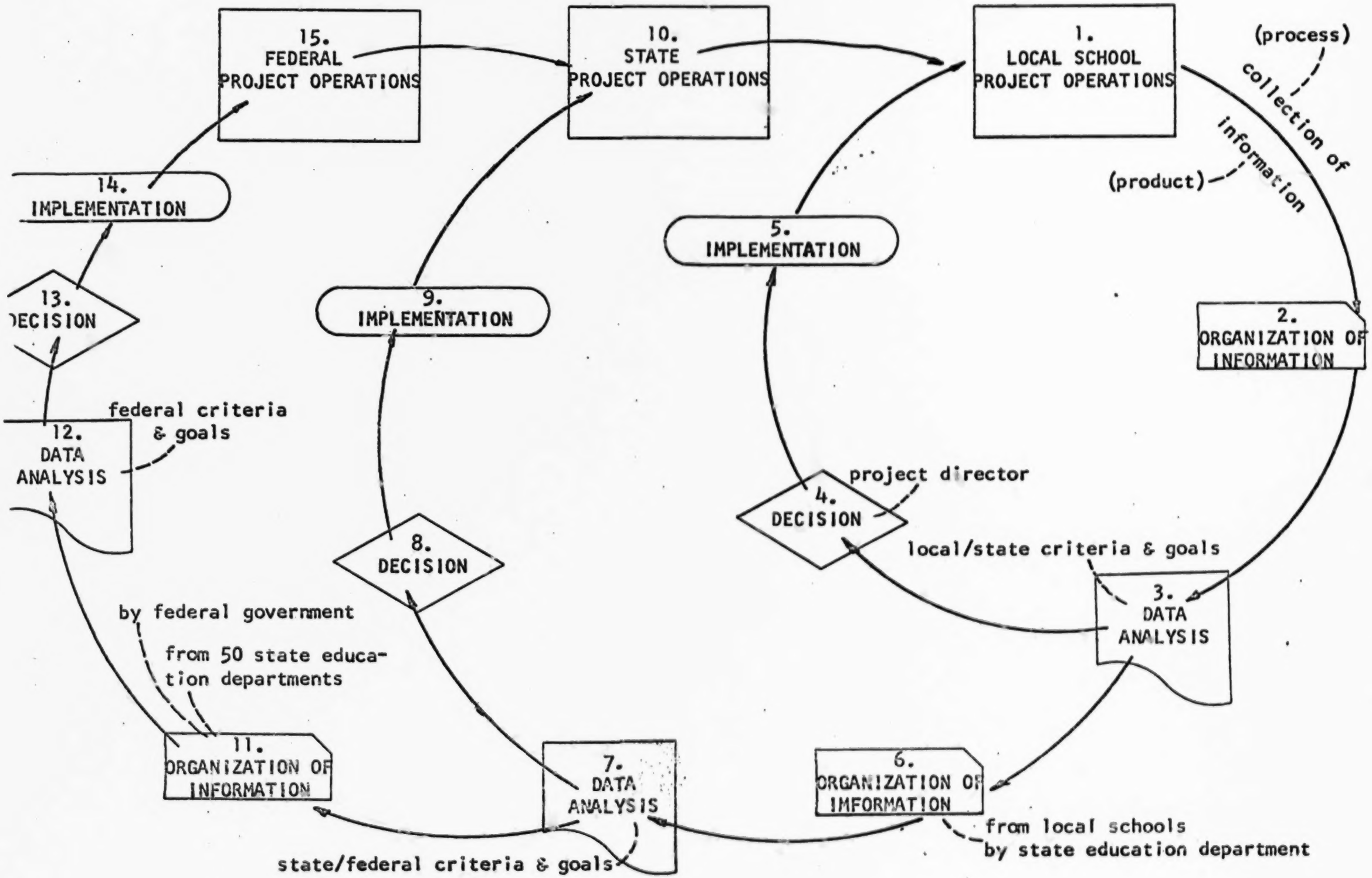
Please direct your attention to Figure 2. This contains a feedback control loop which is intended to delineate the evaluation loops depicted in Figure 1. The specific intent is to clarify the decision function required by Title I. The inner loop depicts the local school activities; the intermediate loop, the state activities; and the outer loop, the Federal activities.

Starting at block 1, we have the initial local school project operations. Proceeding to the right, data collection is depicted by the first segment of the curved line. This occurs at the local school level and is seen to be purely objective. Here, no judgments are indicated or appropriate. What is indicated is the systematic collection of all the data that will be needed for later decisions at local, state, and Federal levels. Thus, it would be useful to know at this stage what kinds of decisions will need to be made at each of these levels.

FIGURE 2

A

FEEDBACK CONTROL LOOP FOR VIEWING THE ROLE OF EVALUATION IN THE TITLE I PROGRAM



It seems to me that basically two kinds of decisions will be required. At the state and Federal levels, the basic decisions will be to increase, decrease, or discontinue entirely the funding of various kinds of projects. At the local level, the most prevalent and significant decisions, in my opinion, will be whether to continue or modify methods and materials as a project proceeds, i.e., as more is learned about an on-going project.

Product Evaluations

For the state and Federal decisions, product or outcome data will be needed. Here, the traditional kind of product evaluation is appropriate. By product evaluation I mean the process of relating outcomes to goals on the basis of specified criteria. This is the kind of evaluation that the majority of the material on evaluation in the Title I guidelines is concerned with. I shall not discuss the details of product evaluation here since they are well-discussed elsewhere. But I do want to make the following assertions about product evaluations as they relate to Title I:

(1) Product evaluations are usually in the form of a comparison of pre-project and post-project performance.

(2) The best that such evaluations can do is to indicate whether the project outcomes were higher, the same as, or lower than some standard.

(3) Product evaluations provide the kinds of information that will be needed for decision-making at the state and Federal levels, but are not adequate for assisting in the decision process at the local level.

(4) To design effective product evaluations, it is necessary to know what the objectives and associated criteria are for the Federal, state, and local levels.

(5) For the evaluative data to be comparable for similar projects, it will be necessary that product evaluation plans be specified at both the Federal and state levels.

As noted, product evaluations do not provide the information that is needed for making decisions at the local level. Product evaluative data usually will be available only annually, and then will contain only the gross information about the effectiveness of an overall project. Such data will be available too infrequently and will lack the specificity required for adequate project control and direction at the local level.

Process Evaluations

For local control and decisions, I believe that a process type of evaluation is required. Process evaluation, as I define it, is the method of relating process to product. The overall strategy is to identify and monitor, on a continuous, molar, non-interventionist basis, the potential sources of failure in the project, e.g., people,

facilities, equipment, and materials. Roger Barker[4] has recently characterized this approach as transducer type inquiry in an ecological context. Here, the evaluator does not exercise experimental control over the situation, nor does he manipulate the situation in any way. Rather, he accepts the situation as it is and as it evolves under real life circumstances. He then "bugs" the situation as best he can by focusing his best observation and other non-interventionist data collection techniques on those aspects of the project which are most crucial to its success. The nature of such an evaluation is multivariate, and not all of the important variates can be specified prior to the initiation of a project. The process evaluator consequently focuses on those variates which are theoretically important, but also remains alert to any unanticipated, but significant events that may occur along the way. In summary, under process evaluations, data are collected day-by-day, organized systematically, analyzed periodically (e.g., weekly), and reported as often as the project director may require such information (e.g., monthly).

I should now like to direct your attention to block number 2 of Figure 2: the organization of data. Here, data would be coded, processed (e.g., keypunched) and filed. This would be necessarily a day-by-day process.

---

[4] Roger G. Barker, "Explorations in Ecological Psychology," American Psychologist, 20 (1965), 1-14.

As the state and local reports were required, the data, both process and product collected at block 2, would be analyzed (at block 3) in accordance with the requirements of these reports. The process data would be analyzed frequently to determine whether the potential sources of failure were becoming actual failures, whether the procedures were following the project design, whether the projects were on schedule, and whether the design or procedures should be modified. Once the outcomes were determined, the process data could also be used to diagnose the outcomes. The product data, on the other hand, would be analyzed infrequently. Essentially, this type of evaluation would determine the overall effectiveness of the project in meeting the prescribed, general goals.

The evaluator would send process evaluation reports to the project director quite frequently. On the basis of these reports, the project director would make decisions (at block 4) about a project. These decisions would be implemented (at block 5) and, as a consequence, an on-going project would undergo frequent modification (block 1). As I view it, this cycle is continuous, and the frequency of process reports is limited only by the frequency with which they are required by the project director. The purpose of these process reports is obviously to assist in project control and evolution.

Returning to block 3, product evaluations would be prepared at least annually and sent to the state education department. The state education department would, of course, receive such reports from all public school districts in the state. The immediate task of the state

education department would, therefore, be to organize the reports into types of projects and to combine the data for similar projects (block 6). Next, this organized data would be analyzed (at block 7) to determine the strengths and weaknesses of the state-wide program. Based on these results, the state coordinator for Title I programs would make decisions (at block 8) about program emphases and state control of the program. These decisions would be implemented (at block 9) and would thus affect the state program operations (block 10). Of course, this would in turn affect the local school project operations (back at block 1).

Returning to block 7, you will note that annual evaluative reports from the fifty states would be sent to the Federal agency. The Federal agency would thus be faced with organizing the data from the fifty states (at block 11) so that major program thrusts could be examined, e.g., those of educational achievement, opportunities, and experiences. These data would then be analyzed (at block 12) on a nationwide basis and the results reported to the Congress and the President. Decisions (block 13) would then be made at the Federal level about program emphases and funding. The implementation (at block 14) of such decisions would affect the operations of the Federal program (block 15). Then the operations of state programs (at block 10) and the local school project (at block 1) would correspondingly be affected.

I have now completed a somewhat rambling tour of Figure 2. Hopefully, I have convinced you of the need for developing and coordinating evaluation plans at Federal, state and local levels. Also, I have tried to indicate that data for all three of these types of evaluations will

be largely collected at the local level, and that these data will form the basis for Federal, state, and local decisions, which will affect local operations. Finally, I have suggested that while product evaluations will suffice for state and Federal decisions, process evaluations are required for effective project control and evolution at the local level. I hope that it is clear that evaluation can be a very positive force for improving education, and that, in fact, educational progress is largely dependent upon effective evaluation. I should now like to move to a discussion of some of the problems associated with fulfilling the role for evaluation that I have just described.

PROBLEMS IN FULFILLING THE ROLE FOR EVALUATION IN THE TITLE I PROGRAM OF E.S.E.A.

In my opinion, there are three basic problems standing in the way of implementing the role for evaluation as outlined above. These are: (1) a lack of evaluation experience in education, (2) a lack of trained evaluators, and (3) a lack of non-interventionist type evaluation designs. I doubt that anyone would refute that these are important evaluation problems. The current question, then, is how to find immediate solutions to these problems, so that evaluation requirements for the Title I program can be effectively met. In Ohio, it was clear that no one kind of agency, i.e., schools, universities, or the state education department, possessed the resources necessary for solving these problems. It was further clear, that these agencies could fair better by working together than by attacking the evaluation problems independently.

In this regard, I am going to proceed by describing the general approach that our Evaluation Center at The Ohio State University is taking in evaluating Title I projects. Then I will close by making several recommendations regarding evaluation.

APPROACH TO EVALUATION OF TITLE I PROJECTS AT THE OHIO STATE UNIVERSITY

The Evaluation Center at The Ohio State University was established last July in response to the needs that were then being posed by the Education Act. The purposes of the Center are: (1) to advance the science of evaluation through extracting general evaluation principles and through developing generalized evaluation designs; and (2) to assist local schools and the Ohio State Education Department in the design and implementation of evaluations. In establishing the Center, it was immediately clear that neither of these purposes could be accomplished unless both were attacked simultaneously. Therefore, our Center is now actively engaged with the Ohio State Department of Education and with several Ohio school districts in developing and implementing evaluation designs. Through this approach, we hope to assist local schools with their immediate evaluation problems and, in so doing, to have the opportunity to study the evaluation process in detail, as it occurs in the <u>real</u> world.

Our basic approach includes a full range of cooperation between the local school and the Center. We assist the local school in writing the evaluation design and the evaluative reports that are required by both the project staff - for making decisions about the on-going project - and by the state education department, for making annual reports to the

Federal government. The local school provides staff members to serve on a school-and-university evaluation team. These staff members are responsible for collecting data throughout the project. The Center takes the responsibility for training the school personnel in evaluation, in organizing and processing data, and in selecting and/or constructing the needed instrumentation.

This approach, in my opinion, has several advantages. Since neither the local school districts nor the universities now possess sufficient numbers of trained evaluators, cooperation of the type described above offers an opportunity to increase the evaluation competencies of local school personnel. Evaluation specialists in universities can provide "on the job" training to these local school personnel. In turn, the local schools can receive expert assistance in conducting the evaluations that are required now and can have increased evaluation expertise on their own staffs after the projects are completed. Also such collaboration will afford the universities an opportunity to study the evaluation process in the context of ordinary circumstances. This should facilitate the production of generalized evaluation designs which meet the requirements of school situations.

RECOMMENDATIONS

I would now like to conclude this presentation by listing several recommendations for your consideration. I realize that these are untried, and in that respect flimsy. However, they may provide a handle for some of the tough problems with which we shall be dealing during the remainder of this conference. If so, more than half of my purpose this morning will have been achieved. Here, then, are my recommendations.

1. Specific state plans for evaluation should be established. These should be set up immediately, and they should be sufficiently specific so that comparable data can be collected from similar projects. Statewide objectives and criteria should be determined. Instruments should be selected and/or developed for measuring these objectives. Projects should be identified and classified according to these objectives. Statewide pretest norms on these instruments should be developed. Finally, these instruments should be administered on a statewide basis. By such a method, the effectiveness of the Title I program in the state could be assessed. Without a specific plan, however, there will be little likelihood of making any sense out of the statewide picture. Unless my current information is faulty, at most only a few states have thus far prepared such plans.

One major objection to this suggestion is that statewide testings would be too time consuming. However, recent developments in measurement theory by Frederic Lord[5] and others have shown that test norms can be developed with no student attempting more than a sample of the test questions in any one test. This technique would seem to me to have tremendous implications for statewide evaluations of Title I programs. Another more difficult objection is that tests for measuring the many classes of objectives that are possible under the Act do not now exist. However, this leads to my next recommendation.

---

[5]Frederic M. Lord, Estimating Norms By Item-Sampling, Educational and Psychological Measurement, Vol. XXII, No. 2, 1962

2. Universities, state education departments, and local school districts should collaborate to provide effective evaluations. Such collaboration is obviously indicated if specific statewide plans are to be implemented. Direction must come from the state education departments. Universities can assist by developing the needed instrumentation, suggesting appropriate data analysis designs, and conducting computer data analyses. Of course, the local schools will plan and implement the basic data collections and analyses. It seems to me that all of these functions can be accomplished better in concert than independently.

3. The evaluation design should not constrain or interfere with the project design; rather, the evaluation design should contribute to the continual improvement of the project. While project designs will initially be based on the best knowledge that is available to the project director, these designs should be amenable to change and improvement as more is learned about the operations of the projects. It would therefore seem inappropriate to employ rigorously controlled experimental designs for evaluating projects. Such designs usually require that a narrow, constant treatment effect be applied in an error-free, laboratory-like context. In the biological sciences, the importance of an error-free context for experiments is seen in the great lengths that are taken to produce and employ gnotobiotic, or germ free, situations. In the social sciences, we see the same concern over error variance in the complicated statistical designs which have been developed in order to extract as much of the error variance as possible. It is patent that Title I-type projects will not be developed in, or installed into, anything like error-free situations. It will also not be possible

to hold constant or to account for all of the interacting variables through the use of statistical designs that are presently available. (Indeed, one of the jobs for evaluation is to determine what the important variables are.) Further projects ultimately will have to be designed to cope with the circumstances of ordinary school operations. Therefore, it would seem appropriate to employ evaluation designs which are non-interventionist, and which periodically provide objective data upon which decisions about modifying the project design and procedures can be based.

4. Both product and process evaluation designs are needed. The product evaluations would be used to determine the overall effectiveness of the project, i.e., the extent to which specified standards have been reached. Such evaluations would provide the outcome data that will be required by the state and Federal governmental agencies. Probably the best non-interventionist types of product evaluation designs are the quasi-experimental designs that have been suggested by Campbell and Stanley and others.[6,7]

The process evaluations would provide information useful to the project staff for (a) detecting any deviation of the actual project from the project design, (b) improving the on-going project, (c) identifying possible problem areas in advance, (d) determining the strengths and weaknesses of the project design and procedures, (e) refining and

---

[6]Donald T. Campbell and Julian C. Stanley, "Experimental and Quasi-Experimental Designs for Research on Teaching," N. L. Gage, editor, Handbook of Research on Teaching, Ch. 3, Pp. 171-246.

[7]Gene V. Glass, "Evaluating Testing, Maturation, and Treatment Effects in a Pretest-Posttest Quasi-Experimental Design, American Educational Research Journal, Vol. 2, March 1965, Pp. 83-89.

expanding the project design, (f) coordinating the over-all activities of the project schools, (g) facilitating communication among the project personnel, (h) diagnosing expected and unexpected outcomes of the project, and (i) communicating characteristics of the project to interested parties. The basic design for conducting process evaluations would be to employ resident evaluation specialists. These persons would be responsible for collecting process data on a day-to-day basis. Barker[8] has recently described this approach. The process evaluator would use such techniques as interviews, classroom interaction analysis, PERT (Program Evaluation and Review Technique), Q-sort instruments, questionnaires, socio-drama, observation, and diaries. Medley and Mitzel[9] have summarized and discussed a number of such techniques.

5. Data collections should be objective; therefore, it would seem appropriate to employ a special evaluation staff which is independent of the project staff. At the level of data collection, information should be collected as objectively as possible, and the information should not be tempered by the judgments of the evaluator toward that which is being evaluated. The information should essentially be raw, untampered-with data. Judgments, or evaluation, would seem to be appropriate only after a full set of data had been pulled together.

This requirement would not seem to present a problem in the area of product evaluations. Such evaluations will essentially employ objective-

---

[8]Roger G. Barker, "Explorations in Ecological Psychology," American Psychologist, 20 (1965), 1-14.

[9]Donald M. Medley and Harold E. Mitzel, "Measuring Classroom Behavior by Systematic Observation," N. L. Gage, editor, Handbook of Research on Teaching, Ch. 6, Pp. 247-329.

type instruments, such as standardized tests. A problem would seem to exist, however, in the case of process evaluations where we will have to contend with the potential problem of human bias, since the essential instrument for obtaining process data is the resident evaluator. It would thus seem desirable that the resident evaluator not have vested interests in the project, and that he be trained in the effective employment of observational data collection techniques. For these reasons, I would suggest that the process evaluations should not be conducted by the project staff, that instead a special evaluation staff be set up, trained, and employed to do this job. It would be desirable, however, that the members of this evaluation staff be drawn from the local school system. In this way the school could increase its evaluation competencies, through what I have described above as "on the job training and experience," for their staff.

6. Schools should invest adequate resources into the design and implementation of evaluations, if these are to become effective instruments for improving education. Title I projects will vary so widely, and my own experience is so limited, that I will not attempt to recommend dollar amounts for conducting evaluations. I do feel certain, however, that it will be necessary to invest more resources into evaluation than we have ever done in the past. To invest less than is required for producing sound, usable data would not seem wise. If there is one thing that educators can learn from industry, it is that wise investments in evaluation can pay off in progress.

CLOSING

In closing, I certainly have no illusions that I have solved any of the problems for which we are convened. I welcome your comments and criticisms and look forward to our association during the remainder of the conference. It has been a real pleasure to be here this morning, and I thank you for your attention.